Validating an office simulation model using RFID technology

V. Tabak, B. de Vries, J. Dijkstra
Eindhoven University of Technology
Faculty of Architecture, Building, and Planning
Design Systems group
PO Box 513
5600 MB Eindhoven
The Netherlands
V.Tabak@bwk.tue.nl

Key words: Office utilisation simulation, building simulation system, validation, RFID, space utilisation.

Abstract: This paper presents the validation of an office utilisation model for the research project called "User Simulation of Space Utilisation (USSU)". The result of this research is a system that can be used for analysing and evaluating the space utilisation of a building for any given organisation. A system for building usage simulation that produces data about activities of members of an organisation can substantially improve the relevance and performance of building simulation tools. This is relevant for engineering domains as well as for architects to evaluate the performance of a building design. For a thorough evaluation of the system an experiment was executed for assessing its predictive quality in the context of a real building, organisation and actual human behaviour; this experiment was executed using RFID technology. The result of the experiment was observed data about the space utilisation of the selected organisation. These data were compared with the space utilisation predicted by the USSU system to evaluate the simulation model. The validation of USSU showed that there were no significant differences between the predicated and observed activity behaviour. As a consequence, the output of USSU is considered to be valid.

1.   INTRODUCTION

Activity and location schedules are input for office utilisation simulations. These schedules however, are often assumptions rather than based on measured observations and resulting descriptive and predicting models. Thus, the results of such simulation systems are tentative at best and may often be misleading. Therefore, a more advanced scheduling method is needed that adequately represents real-life complexity of human activity and location schedules. The main objective of research project User Simulation of Space Utilisation (USSU) was to develop a system that can be applied for analysing and evaluating the space utilisation of a building for any given organisation. The system generates activity schedules that provide a representation of human activities that are executed in building spaces. An activity schedule not only describes which activities are performed and at which location, but also the route that is followed between the locations of these activities. These activity schedules are a source of dynamic input data for building usage simulation tools. Reliable data on human movement is scarce. It is valuable input for several research areas. For instance, the relevance and performance of building simulation tools like indoor climate simulations or working conditions assessments will substantially improve when realistic input data is applied. If reliable human movement models can be created, then these models can not only be used to analyse existing situations, but also to simulate new building designs taking the digital design as input. This is also relevant for architects to evaluate the performance of a building design.

USSU models the scheduling behaviour of employees following their heuristics. In other words, the scheduling method mimics the behaviour of real human beings when scheduling activities (Tabak, de Vries, et al., 2006). The proposed scheduling method is not producing schedules for planning purposes based on an optimisation method. The goal of the scheduling method is not to find the optimum activity schedule for instance with regard to the priority of its activities.

USSU serves as a pre-processor for other simulation systems that need real-life data about the location of persons at a specific time, such as indoor climate calculations, evacuation simulations and working condition assessment simulations. The underlying model integrates activity modelling and workflow modelling constrained by spatial conditions of a building (Tabak, de Vries, et al., 2004). The combination of these two methods provides a solid basis that ensures process consistency, inclusion of individual (re)scheduling behaviour and allows execution in compressed time intervals. With this model, space utilisation can be simulated at a high level of realism instead of make assumptions as usual in many building simulation systems. This is possible through the incorporation of human

activities as they are executed in real life, split up into skeleton activities (i.e. workflow dependent activities) and intermediate activities (i.e. social and/or physiological activities). Scheduling of activities is executed under bounded rationality. Persons are only aware of their own agenda which evolves in compressed time.

The scheduling method can accommodate real-life complex organisational structures and maintains consistency between the activity schedules. Most existing approaches to model activity schedules assume an individual-based decision making process. These models do not take into account the interaction between individuals. A special feature of the USSU system is the incorporation of interactions between activities (e.g. meetings). The interaction between activities contributes significantly to the realism of the activity schedules.

The final step in this project was to validate the system. For a thorough evaluation of the system an experiment was performed using RFID technology. The goal of this experiment was to collect data for assessing the predictive quality of USSU in the context of a real building, organisation and actual human behaviour. The experiment resulted in observed data about the actual space utilisation of a chosen organisation. The observed data were compared with the space utilisation predicted by USSU to evaluate its predictive quality. This paper focuses on the validation approach and observation method selected with regard to the validation of USSU.

This paper is organised as follows. First, the test case chosen for the validation of USSU is introduced. After this, the validation approach is unfolded; this section explains how the actual validation was performed and which statistical tests were used to asses the goodness-of-fit. Next, the observation method is described; this section will treat the setup of the experiment to collect the real human behaviour. Then, the validation results are treated. Finally, the last section reflects on the validation of USSU.

2. TEST CASE DESCRIPTION

For the validation process of USSU a specific organisation and a building had to be selected. In order to be able to capture the space utilisation of an organisation in the real world, the decision was made to choose an existing organisation already housed in an office building. At first private companies were asked to participate. However, the contacted private companies were reluctant to participate because of privacy reasons ('Big brother is watching you') and because participating required some effort on their side. As a result of these issues and due to the fact that the required input data was readily available, the decision was made to perform the required experiments on the ninth floor of the building of the faculty of Architecture, Building,

and Planning; the floor where the Design Systems group is housed (see *Figure 1*). The second chair located on this floor, Structural Design, was willing to participate in this experiment.

Figure 1. Simulated office environment of the test floor.

3.   VALIDATION APPROACH

Validation in the context of this project related to determining the validity of the output of the USSU system. In other words, how accurately can the prototype predict the space utilisation of an organisation? According to Kleijnen (1995):

"*Validation, however, can not be assumed to result in a perfect model, since the perfect model would be the real system itself (by definition, any model is a simplification of reality). The model should be 'good enough', which depends on the goal of the model.*".

In order to validate a simulation model such as USSU, output of this model is compared with observations performed in the real world (see *Figure 2*). However, before the USSU system can be applied for creating the activity schedules, the system had to be calibrated. In order to obtain reliable results the input data needs to be constructed based on the context in which the validation is performed. For predicting the activity schedules the system requires detailed input (Tabak, de Vries, et al., 2006).
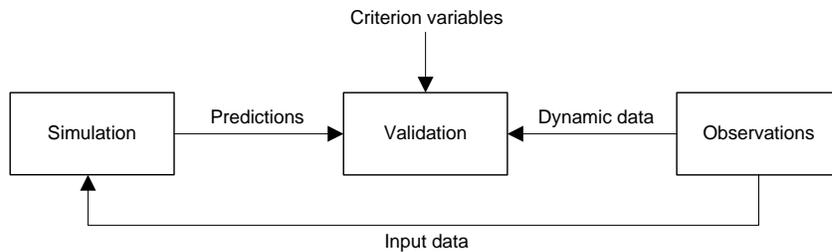
Figure 2. Overview of the elements of the validation process (Daamen, 2004).

The validity of USSU was determined by comparing predicted space utilisation with observed space utilisation. Activity schedules predicted by the USSU prototype were compared with activity schedules observed in the real world. Before the validation process could be started two essential questions had to be answered:

- How should the actual validation be performed?
  In which way should the observed and predicted activity schedules be compared with each other?
- When is the (output of the) prototype considered to be valid?
  When is the predicted space utilisation a sufficient accurate representation of the reality (i.e. observed space utilisation)?

These questions are answered in the following two sections.

3.1   How Should The Actual Validation Be Performed?

In the real world the activity schedule of an employee differs per day, in terms of start/end time of the schedule, the order of activity, their timing and last but not least the types of activities. In other words, per working day an employee performs a different set of activities in varying sequences. As a result the observed activity schedules differ per day. Due to variation in the activity scheduling process implemented in the prototype also the predicted activity schedules vary per day. So both the predicted and observed activity schedules vary per day. This makes comparing these two types of activity schedules a challenge. A direct and qualitative comparison of predicted and observed activity schedules on among others the order and timing of activities is not relevant; at least not determining the validity of the system. It is without any doubt that on these points the activity schedules will vary. For validating USSU it is more relevant to analyse its output on a higher, more abstract level. The USSU system could (eventually) be applied to generate input for building performance simulation systems, like evacuation simulations or working conditions assessments. In general, this type of building simulation tools analyses the performance of a building based on a set of performance indicators, like the walking time per individual, number

of persons per space in time and the usage of facilities. The values of these performance indicators can be deduced from the predicted activity schedules. The same indicators which are used for evaluating a building can also be used for validating USSU. To assess the validity of the system the predicted space utilisation was compared with the observed space utilisation on a set of performance indicators, so-called criterion variables (see *Figure 2*). These criterion variables (e.g. the usage of facilities or movement behaviour of employees) specify the aspects on which the comparison of the observed and predicted activity schedules was performed. The values of the criterion variables were derived from both the predicted and observed activity behaviour of all employees.

The following criterion variables were chosen for validating USSU:

- Zone utilisation
- Mean walking distance of employees.

These criterion variables are discussed below.

3.1.1   Zone Utilisation

The first criterion variable relates to the amount of time a zone is used during a working day, either observed in the real world or simulated using USSU. The zone utilisation percentage is a proportion of the total simulated or observed time. It is calculated by dividing the total time a zone is used by the total simulated or observed time. In formula:

$$P_x = \frac{T_x^{zone}}{T_s} \, , \; T_x^{zone} = \sum_{i=1}^{n_x} \Delta T_{a_{x,i}} \qquad\qquad (1 \,\&\, 2)$$

where:

P_x is utilisation of zone x for an average working day

T_x^{zone} is total time zone x is used during simulation/observation

T_s is total simulated or observed time

$a_{x,i}$ is activity a_i using zone x

$\Delta T_{a_{x,i}}$ is duration of activity $a_{x,i}$

n_x is total number of activities using zone x

3.1.2   Mean Walking Distance of Employees

The second criterion variable examines the movement behaviour of employees. The activity schedule of an employee reveals the required

movement behaviour as a result of its activities. First the total walking distance for all simulated or observed days is calculated as follows:

$$W_e^{total} = \sum_{i=1}^{m} W_e^{day}, \; W_e^{day} = \sum_{i=1}^{n-1} L_{[a_{e,i}, a_{e,i+1}]} \qquad (3 \, \& \, 4)$$

where:

W_e^{total} is total walking distance of employee e for all simulated/observed work days

W_e^{day} is total walking distance of employee e for one simulated/observed work day

m is total number of simulated or observed working days

$L_{[a_{e,i}, a_{e,i+1}]}$ is length of route between location of activity $a_{e,i}$ and location of activity $a_{e,i+1}$

n is number of activities in the activity schedule of employee e for a work day

Next, the mean walking distance of an employee on a working day is calculated:

$$W_e^{mean} = \frac{W_e^{total}}{m} \qquad (5)$$

where:

W_e^{mean} is mean walking distance of employee e for a working day

3.2   When Is The Prototype Considered To Be Valid?

The above mentioned validation approach resulted in an observed and predicted set of data for each criterion variable. These two sets of data were compared with each other to determine the predictive quality of USSU. In other words, the two sets were used to assess the goodness-of-fit between the predicted and observed activity behaviour. The goodness-of-fit was measured using a combination of the following two tests, namely:

- Student's *t*-test combined with the correlation coefficient.
- Variability test.

The *t*-test is a statistical hypothesis test in which the statistic has a Student's *t*-distribution if the null hypothesis is true. Given two data sets,

each characterised by its mean, standard deviation and data samples, the *t*-test can be used to determine whether the means are distinct, provided that the underlying distributions can be assumed to be normal. If the calculated *t*-value is above the threshold chosen for statistical significance (usually the 0.05 level; α=0.05), than the null hypothesis that the two groups do not differ is rejected in favour of an alternative hypothesis, which means that the groups differ. Once a *t*-value is calculated, the tail-end probability, called *P*-value, can be found using a table of values from Student's *t*-distribution or using the underlying calculation formula of the value. If the *P*-value is below the chosen threshold (i.e. *P* < 0.05) than the null hypothesis is rejected.

There are different versions of *t*-test depending on whether the samples are independent of each other or paired, meaning that each member of one sample has a relationship with a member of the other sample. The paired samples *t*-test can only be applied when the corresponding correlation coefficients proved to be significant; otherwise the independent samples *t*-test is applied. The observed and predicted data sets were compared on one criterion variable at a time. The closer the *P*-value comes to 1 the smaller the differences between the observed and predicted data set values.

Although, the variability test is not a standard statistical test it gives further insight in the differences between the observed and predicted sets of data. The smaller the variability is (i.e. the closer it gets to zero) the better the match between the two data sets, e.g. the observed and predicted activity behaviour. The variability can be calculated using the following formula:

$$v^2 = \frac{\sum_{i=1}^{n}(s_{c,i} - o_{c,i})^2}{n} \qquad (6)$$

where:

v^2	is variability
$s_{c,i}$	is predicted value i of the criterion variable c
$o_{c,i}$	is observed value i of the criterion variable c
n	is the number of observations

4.   OBSERVATION METHOD

There are many possible ways of collecting data about the behaviour of people, such as through paper/digital activity diaries, by using technical equipment (e.g. video camera's or infrared sensors) or by discreet shadowing of people (Teknomo, Takeyama, et al., 2001; Kerridge, Hine, et al., 2005,

Tan, 2003, Arentze, Hofman, et al., 1997). Each method has its strong and weak points. Collecting data about human activity behaviour using paper/digital questionnaires is rather obtrusive and puts a relative high demand on the participants; they have to record all their activities by themselves. There is a realistic chance that people forget to enter activities, albeit not deliberate (Ettema, 1996). The use of technical systems (e.g. video cameras) to record the movements of people is expensive (e.g. it requires a lot of cameras to cover a whole floor and to collect data for a long period requires a lot of data storage). These systems also require a lot of post processing time (e.g. the data recorded using separate cameras have to be combined in order to track individuals across the whole floor). In addition, such systems are highly sensitive in the context of privacy issues, especially in the case of video cameras. Furthermore, several sources in literature (Sundstrom, 1986; Martin and Bateson, 1993) suggest that the knowledge of being observed influences the behaviour of people (the so-called Hawthorne effect). The ideal situation for collecting data about human movement would be when the participants do not know that they are being observed (Fatah, Schieck, et al., 2006). However, in reality this is hardly feasible, not only for operational reasons (i.e. how to track individuals across a space if they are not 'equipped' with a certain traceable feature), but also because of privacy reasons.

A relative new technology called RFID (Radio Frequency Identification) could be the key for a non obtrusive way of collecting data about human movement and activity behaviour. Using this system, participants only have to carry a small device and the RFID system automatically registers their movements. During the observation period participants themselves do not have to perform any additional actions besides performing their normal activity behaviour. Due to the passive nature of this system the behaviour of people is only influenced in a limited manner and probably only the most in the beginning of the experiment. This system results in observed data about the daily movement behaviour with a relative high degree of precision.
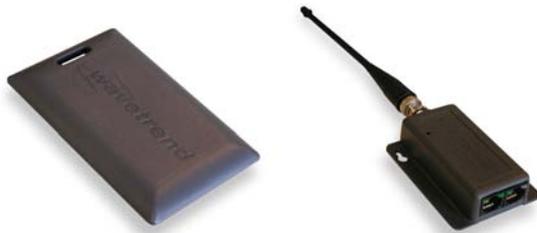

Figure 3. RFID tag and reader.

4.1   RFID

An RFID system consists of readers and tags (see *Figure 3*). An RFID tag is a device which can be remotely accessed to retrieve data stored in its chip. Each tag is equipped with an antenna to receive and respond to radio-frequency queries from an RFID reader. There are two types of RFID tags: passive and active tags. Passive tags do not have an internal power supply; the power required to transmit a response is induced through the incoming radio-frequency signal. These tags can only be read out from short distances (up to several meters). However the active tags can have a read distance of up to several hundreds of meters. Therefore these tags are equipped with an internal power source. An active RFID tag will send out a signal for example every 1.5 seconds. The battery of these tags last up to 5 years. In this experiment the active tags were applied in combination with a number of readers. Currently, RFID technology is used by some Dutch organisations for access control and as a means of working hours registration, but up to now never on the detailed level as it was used in this experiment.

All employees working on the test floor were asked to participate in the RFID experiment. Roughly 50 employees are officially located on this floor. In the end 37 persons accepted the request. Each person was asked to wear a RFID card for a period of 3 months. On the office floor 16 receivers were installed. The placement of readers was such that the real movement behaviour of the employees could be tracked throughout the whole floor.

In 46 days the RFID system recorded about 360,000 events; each event relates to one of the 37 participants entering or leaving one of the RFID zones. These events had a total duration of more than 32,000 hour.

4.2   Post Processing

When the data collected in the RFID experiment is compared with the predicted space utilisation attention had to be paid to the way in which the activity behaviour is collected. Although the USSU system predicts space utilisation on the level of (parts of) building spaces, in the RFID experiment the space utilisation was observed on level of sets of spaces, called zones (*Figure 4*). The main reason for this was that it proved to be too costly to equip each space with an RFID reader. To be able to compare the observed space utilisation with the predicted space utilisation, the predicted data had to be aggregated to allow for comparison and hence model validation.

■ RFID reader   ⊡ RFID reader on upper floor   ⬜/▤ RFID zones

Figure 4. RFID zone layout.

4.3   Problems Related To RFID Setup

As mentioned above the floor was divided into zones; each zone was equipped with 1 or 2 readers to detect the presence of a tag (and so an employee). A reader uses a spherical radio field to detect tags. When two or more readers are nearby each other, the spherical fields of some readers can overlap; the distance at which this occurs depends on how the reader was set up (e.g. the threshold of the signal strength used by a reader to detect a tag and the type of antenna equipped with the reader). This means that a tag can be in 2 or more zones at the same time. The tracking software selects the reader (and so the zone) which has the highest signal strength. However, as the signal strength of a tag recorded by the reader is influenced by distortions and reflections, the signal strength is constantly changing. During the experiment this caused tags to jump from one zone to another when in fact they were not moving. This was taken into account when validating USSU.

5.   VALIDATION RESULTS

This section discusses the results of the validation. First the results of the validation with regard to the first criterion variable (i.e. zones) are discussed. Then, the validation results related to the second criterion variable (i.e. movement behaviour of employees) are treated.

5.1   Zone Utilisation

In the RFID experiment human activity behaviour was observed on level of sets of spaces, called zones. The problems with the RFID setup had only a minor influence on the zone utilisation and were consequently disregarded.

Table 1 indicates a strong correlation between the predicted usage of zones (USSU) and the observed usage of zones (RFID); this result is significant at 0.05 level. As a consequence, it was allowed to perform a paired samples *t*-test. The paired samples *t*-test for the zone utilisation shows a powerful outcome (*Table 1*). Consequently, the null hypothesis could not be rejected. In other words, there were no significant differences between the predicted (USSU) and observed (RFID) usage of zones. This rather strong outcome of the *t*-test was further emphasised by the variability test. The results of this test also indicate negligible differences between the USSU and RFID data sets with regard to the utilisation of zones.

Table 1. Validation results for the usage of zones.

Criterion variable	Correlation coefficient		Paired-samples *t*-test		Variability	
	r-value	*P*	*t*-value	*P*	value	std. dev.
Zone utilisation	0.77	0.01	0.00	1.00	0.00	0.00

5.2   Employees

Not only did the RFID tracking system store the amount of time an individual was in a certain zone, it also stored the routes which were followed between the RFID zones. This meant that the data collected in the RFID experiment could also be used to validate USSU on the level of human movement behaviour, in particular the mean walking distance per employee. However, due to above mentioned problems related to the setup of the experiment in conjunction with the RFID technology, a large part of the stored activity data was not usable for this comparison. As a result of the overlapping RFID zones, tags (i.e. people) jumped from one zone to another when they were not moving. This meant that the walking distances for individuals with a workplace in one of these zones were highly unreliable. Therefore these individuals were removed from the RFID data set and were not taken into account for the validation of USSU with regard to the movement behaviour of employees.

Table 2 shows the correlation coefficient between the USSU and RFID data sets in relation to the mean walking distance. The correlation coefficient was significant and indicated a strong link between the two data sets. The results of the paired samples *t*-test for the mean walking distance were significant (*Table 2*). This meant that for this criterion variable the null

hypothesis could not be rejected. With regard to the mean walking distance there were no significant differences between the two data sets. Finally, the variability test also suggests only minor differences between the two data sets with regard to movement behaviour (*Table 2*).

Table 2. Validation results for movement behaviour of employees.

Criterion variable	Correlation coefficient		Paired-samples *t*-test		Variability	
	r-value	*P*	*t*-value	*P*	value	std. dev.
Time percentage	0.70	0.04	1.62	0.14	0.05	0.08

6. DISCUSSION

The validation of USSU showed that there were no significant differences between the predicted and observed activity behaviour. This is further emphasised by that fact that all statistical tests (i.e. Student's *t*-test, correlation coefficient and variability test) supported each other and pointed in the same direction. As a consequence, the output of USSU is considered to be valid. In other words, USSU can be used to accurately predict the space utilisation of an organisation (at least when applied within the limitations set for this project).

RFID technology allows for a non obtrusive way of collecting data about human movement. Participants of the experiment only had to carry a small RFID tag, for instance in their wallet, and the RFID system automatically registered their movements using a number of strategically placed readers. The RFID system made it possible to track the movements of all participants across the floor and thereby collecting data about the real movement behaviour of the participants. To improve the accuracy of the collected data (e.g. to 'filter' out infrequent behaviour) the human movement behaviour had to be collected for a relative long continuous period. Therefore and as the RFID system did not put a high demand on the participants, it was decided to collect the data for a considerable period of time, namely 3 months. As a result, the observed human activity behaviour was averaged over a long time interval. In spite of problems related to the RFID setup (i.e. overlapping zones which caused tags to jump for one location to another) the experiment resulted in realistic data about the daily movement behaviour.

A disadvantage of the RFID experiment is that it only delivers data about the utilisation of zones and the movement of employees between these zones. In order to thoroughly validate USSU data was needed on the level of types of activities participants are performing. Due to the required level of detail of activity data the most logical method of obtaining of these data is through activity diaries, which are maintained by the participants themselves

during a working day. Other methods, like questionnaires or video recording, do not reveal all the required data. By administering the activity diaries during the day instead of at the end of the day or the next day the chance of forgetting activities by respondents is small.

A second experiment was performed in which only direct colleagues belonging to the Design Systems group were asked to maintain a diary of their daily activities. To limit the load on the participants, this experiment was performed parallel to the RFID experiment and had the same duration, namely 3 months. Entering data with regard to the daily activity diaries was made as simple, user-friendly and intuitive as possible by applying a purpose-built survey system. Any diary based system works post hoc, meaning that test persons may enter erroneous information because of problems of recall. Our system is no exceptions to this, but we hope to have minimised the risk by making the system as accessible and easy to use as possible. In the near future, we present the results of the second experiment for validating USSU.

7.   REFERENCES

Arentze, T., F. Hofman, N. Kalfs and H. J. P.Timmermans, 1997, "Data needs, data collection and data quality requirements of activity-based transport demand models", in: *International Conference on Transport Survey Quality and Innovation*, Grainau, Germany.

Daamen, W., 2004, *Modelling passenger flows in public transport facilities*, DUP Science, Delft, The Netherlands.

Ettema, D.F., 1996, *Activity-based travel demand modelling*, Technische Universiteit Eindhoven, Eindhoven, The Netherlands. PhD thesis.

Fatah gen. Schieck, A. Penn, V. Kostakos, E. O'Neill, T. Kindberg, D. Stanton Fraser and T. Jones, 2006, "Design tools for pervasive computing in urban environments", in: *8th International conference on design and decision support systems in architecture and urban planning*, p. 467-486, Springer, Dordrecht, The Netherlands.

Kerridge, J., J. Hine and M. Wigan, 2001, "Agent-based modelling of pedestrian movements: the questions that need to be asked and answered", *Environment and planning B*, 28, p. 327-341.

Kleijnen, J.P.C., 1995, "Verification and validation of simulation models", *European journal of operational research*, 82 (1), p. 145-162.

Martin, P. and P. Bateson, 1993, *Measuring behaviour: An introductory guide*, Cambridge University Press, Cambridge, UK.

Sundstrom, E., 1986, *Work places: the psychology of the physical environment in offices and factories*, Cambridge University Press, Cambridge, UK.

Tabak, V., B. de Vries and J. Dijkstra, 2004, "User Behaviour Modelling - Applied in the context of space utilisation", in: J.P. van Leeuwen and H.J.P. Timmermans (eds.) *Developments in Design & Decision Support Systems in Architecture and Urban Planning*, Eindhoven, Netherlands: 141-156.

Tabak, V., B. de Vries, J. Dijkstra and A.J. Jessurun, 2006, "Interaction in Activity Location Scheduling", in: *Proceedings of the 11th International Conference on Travel Behavior*

Research, Kyoto, Japan. CD-rom.

Tan, A.A.W., 2003, *The reliability and validity of interactive virtual reality computer experiments*, Technische Universiteit Eindhoven, Eindhoven, The Netherlands. PhD thesis.

Teknomo, K., Y. Takeyama and H. Inamura, 2001 "Tracking system to automate data collection of microscopic pedestrian traffic flow", in*: Proceedings of The 4th Eastern Asia society For transportation studies*, 3 (1), Hanoi, Vietnam,. p. 11-25.